Bezpečnosť asymetrického šifrovania

Martin Stanek

2025

KI FMFI UK Bratislava

Úvod

- asymetrické šifrovanie
- formálny prístup k definícii bezpečnosti šifrovacích schém
- IND-CPA a IND-CCA2
- dôkazy redukciou na podkladový problém
- ElGamalova schéma
- Fujisakiho-Okamotova transformácia

Šifrovacia schéma

šifrovacia schéma $\Pi = \langle \text{Gen}, \text{Enc}, \text{Dec} \rangle$

- **Gen** – znáhodnený polynomiálny algoritmus
  - vstup: 1^k (k je bezpečnostný parameter schémy)
  - výstup: (pk, sk), kde pk je verejný kľúč a sk je súkromný kľúč
  - priestor správ $\mathcal{M}$ je obvykle určený vygenerovanými kľúčmi
- **Enc** – znáhodnený polynomiálny algoritmus
  - vstup: pk, $m \in \mathcal{M}$
  - výstup: šifrový text $c \leftarrow \text{Enc}_{\text{pk}(m)}$
- **Dec** – deterministický polynomiálny algoritmus
  - vstup: sk, c
  - výstup: otvorený text $m = \text{Dec}_{\text{sk}}(c)$, resp. chybový symbol $\perp$

- korektnosť schémy: $\forall k \in \mathbb{N}; \forall (\text{pk}, \text{sk}) \leftarrow \text{Gen}(1^k)\ \forall m \in \mathcal{M} : \ \text{Dec}_{\text{sk}}\big(\text{Enc}_{\text{pk}}(m)\big) = m$

Ciele a schopnosti útočníka

- definícia bezpečnosti – kombinácia cieľa a schopností útočníka
- najsilnejšia definícia = najslabší cieľ + najsilnejšie schopnosti
- **ciele**:
  - získať súkromný kľúč (úplné rozbitie schémy)
  - získať otvorený text k danému šifrovému textu
  - vypočítať netriviálnu informáciu o otvorenom texte (sémantická bezpečnosť)
  - rozlíšiť správny otvorený text prislúchajúci k šifrovému textu (nerozlíšiteľnosť)
  - modifikovať šifrový text tak, že (neznáme) otvorené texty sú známym spôsobom závislé (nepoddajnosť)
  - skonštruovať korektný šifrový text bez znalosti prislúchajúceho otvoreného textu
- **schopnosti**:
  - znalosť verejného kľúča (teda aj možnosť čokoľvek zašifrovať)
  - prístup k dešifrovaciemu orákulu (potenciálne s rôznymi obmedzeniami)

Definícia bezpečnosti – experiment (IND-CPA)

- výpočtovo obmedzený (polynomiálny) útočník A
- útočník má prístup k verejnému kľúču
- cieľ: rozlíšiť, z ktorého z dvoch otvorených textov vznikol šifrový text

$\mathbf{Ind}^{\mathbf{cpa}}_{A,\Pi}(\boldsymbol{k})$:

$(\mathrm{pk}, \mathrm{sk}) \leftarrow \mathrm{Gen}(1^k)$

$(m_0, m_1, s) \leftarrow A(\mathrm{pk})$, kde $|m_0| = |m_1|$ a s reprezentuje stav výpočtu

$c \leftarrow \mathrm{Enc}_{\mathrm{pk}}(m_b)$, kde $b \in_R \{0, 1\}$

$b_A \leftarrow A(s, c)$

return $b_A \overset{?}{=} b$ (1 ak A určil b správne, inak 0)

Definícia bezpečnosti – IND-CPA

Definícia. Šifrovacia schéma $\Pi = \langle \text{Gen}, \text{Enc}, \text{Dec} \rangle$ je IND-CPA bezpečná, ak pre ľubovoľného PPT útočníka A existuje zanedbateľná funkcia $\text{negl}(\cdot)$ taká, že platí:

$$\Pr\left[\text{Ind}^{\text{cpa}}_{A,\Pi}(k) = 1\right] \leq \frac{1}{2} + \text{negl}(k).$$

– $f : \mathbb{N} \to \mathbb{R}^+_0$ je (asymptoticky) *zanedbateľná*, ak pre ľubovoľný kladný polynóm p:

$$\exists n_0 \in \mathbb{N} \;\; \forall n > n_0 : \quad f(n) < \frac{1}{p(n)}$$

– *nejakú* zanedbateľnú funkciu označíme $\text{negl}(\cdot)$

Poznámky k definícii

- asymptotická definícia
- IND-CPA nie je *najsilnejšia* definícia
- IND-CPA a iné varianty je možné definovať aj pre symetrické šifrovacie schémy
  - CPA: prístup k šifrovaciemu orákulu
  - používané napríklad pri analýze módov blokových šifier
- prečo nemôže byť definícia IND-CPA „nejako perfektná“, teda
  - s výpočtovo neobmedzeným útočníkom $\longrightarrow$ dokáže dešifrovať s pravd. 1
  - s $\Pr[\cdot] = 1/2 \longrightarrow$ aj polynomiálny útočník sa dokáže odchýliť od $1/2$
- deterministické schémy (s deterministickým Enc) nemôžu byť IND-CPA bezpečné
  - A vie šifrovať a výpočtom $\mathrm{Enc}_{\mathrm{pk}}(m_0)$, $\mathrm{Enc}_{\mathrm{pk}}(m_1)$ určí b s pravdepodobnosťou 1
- nerozlíšiteľnosť ekvivalentná pojmu *sémantická bezpečnosť* (nemožnosť vypočítať netriviálnu informáciu o otvorenom texte)

ElGamalova schéma

- $\mathrm{Gen}(1^k)$:
  - vygenerujeme grupu $(G, \cdot)$ prvočíselného rádu q (kde $|q| = k$) a jej generátor g
  - $\mathrm{sk} = x \in_R \mathbb{Z}_q$; $\mathrm{pk} = y$, kde $y = g^x$
  - parametre grupy (G, q, g) sú všeobecne známe, prípadne sú súčasť sk a pk
    - teda napr. $\mathrm{pk} = (G, q, g, y)$
  - priestor správ $\mathcal{M} = G$
- $\mathrm{Enc}_{\mathrm{pk}}(m) = \left(g^t, y^t \cdot m\right)$, kde $t \in_R \mathbb{Z}_q$ (znáhodnené šifrovanie)
- $\mathrm{Dec}_{\mathrm{sk}}(r, s) = s \cdot r^{-x}$
- korektnosť schémy: $s \cdot r^{-x} = y^t \cdot m \cdot r^{-x} = g^{xt} \cdot m \cdot g^{-tx} = m$

Rozhodovací DH problém

- parametre: G, q, g ako v schéme
- problém diskrétneho logaritmu (DLOG):
  - pre dané $y \in_R G$ vypočítať $x \in \mathbb{Z}_q$: $g^x = y$
- výpočtový Diffieho-Hellmanov problém (CDH):
  - pre dané g^a, g^b, kde $a, b \in_R \mathbb{Z}_q$ vypočítať g^{ab}
- rozhodovací Diffieho-Hellmanov problém (DDH):
  - pre danú trojicu prvkov z G rozhodnúť jej *typ*
  - skutočná DH trojica: $\left(g^a, g^b, g^{ab}\right)$, kde $a, b \in_R \mathbb{Z}_q$
  - náhodná trojica: $\left(g^a, g^b, g^c\right)$, kde $a, b, c \in_R \mathbb{Z}_q$

DDH predpoklad

- formalizácia predpokladu, že DDH problém je v danej grupe ťažký
- definujme výhodu útočníka/rozlišovača D:

$$\mathrm{Adv}^{\mathrm{ddh}}_{G,q,g}(D) = \left|\Pr\left[a, b \in_R \mathbb{Z}_q : D\left(g^a, g^b, g^{ab}\right) = 1\right]\right.$$
$$\left. - \Pr\left[a, b, c \in_R \mathbb{Z}_q : D\left(g^a, g^b, g^c\right) = 1\right]\right|$$

- (DDH predpoklad) Pre ľubovoľného PPT útočníka D existuje negl:

$$\mathrm{Adv}^{\mathrm{ddh}}_{G,q,g}(D) \leq \mathrm{negl}(k)$$

IND-CPA bezpečnosť ElGamalovej schémy

Veta. ElGamalova schéma je IND-CPA bezpečná, ak pre príslušnú grupu platí DDH predpoklad.

Dôkaz (1) – modifikovaná schéma Π^*

- nech Π je ElGamalova schéma, A je ľubovoľný PPT útočník
- označme $\varepsilon(k) = \Pr\left[\mathrm{Ind}^{\mathrm{cpa}}_{A,\Pi}(k) = 1\right]$
- vytvorme modifikovanú schému Π^*
  - rovnaký algoritmus Gen
  - upravený $\mathrm{Enc}^*(m) = \left(g^t, g^u \cdot m\right)$, kde $t, u \in_R \mathbb{Z}_q$
  - na dešifrovacom algoritme nezáleží (ani nie je možné dešifrovať)
  - $\mathrm{Ind}^{\mathrm{cpa}}$ experiment závisí len na Gen a Enc
- pozorovania pre šifrový text v Π^*, teda $\left(g^t, g^u \cdot m\right)$:
  - prvý komponent nezávisí na m a je uniformne distribuovaný v G
  - pre ľubovoľné m je druhý komponent uniforme distribuovaný v G (lebo g^u je)
  - teda $g^u \cdot m$ nenesie žiadnu informáciu o m (*one-time pad* nad G)
  - preto $\Pr\left[\mathrm{Ind}^{\mathrm{cpa}}_{A,\Pi^*}(k) = 1\right] = \frac{1}{2}$

Dôkaz (2) – rozlišovač D pre DDH

- skonštruujme PPT rozlišovač D pre DDH problém
  - vstup: G, q, g a trojica prvkov z G, označme ich g_1, g_2, g_3
- D použije útočníka A (bude ho simulovať):
  1. položí $\text{pk} = (G, q, g, g_1)$, teda $y = g_1$
  2. simuluje: $(m_0, m_1, s) \leftarrow A(\text{pk})$
  3. zvolí náhodné $b \in_R \{0, 1\}$ a vytvorí šifrový text $c = (g_2, g_3 \cdot m_b)$
  4. simuluje: $b_A \leftarrow A(s, c)$
  5. return $b_A = b$ (teda 1 pri zhode, 0 inak)

Dôkaz (3) – analýza D

1. na vstupe D je náhodná trojica $\left(g^x, g^t, g^u\right)$

- šifrový text zodpovedá šifrovaniu v schéme Π^*, $c = \left(g^t, g^u \cdot m_b\right)$
- D vráti 1 práve vtedy, keď A uspeje v $\mathrm{Ind}^{\mathrm{cpa}}$ experimente pre Π^*
- teda $\Pr\left[D\left(g^x, g^t, g^u\right) = 1\right] = \Pr\left[\mathrm{Ind}^{\mathrm{cpa}}_{A,\Pi^*}(k) = 1\right] = \frac{1}{2}$

2. na vstupe D je DH trojica $\left(g^x, g^t, g^{xt}\right)$)

- šifrový text zodpovedá šifrovaniu v schéme Π, $c = \left(g^t, g^{xt} \cdot m_b\right)$
- D vráti 1 práve vtedy, keď A uspeje v $\mathrm{Ind}^{\mathrm{cpa}}$ experimente pre Π
- teda $\Pr\left[D\left(g^x, g^t, g^{xt}\right) = 1\right] = \Pr\left[\mathrm{Ind}^{\mathrm{cpa}}_{A,\Pi}(k) = 1\right] = \varepsilon(k)$

Dôkaz (4) – záver

- keďže platí DDH predpoklad rozdiel pravdepodobností oboch prípadov pre D musí byť zanedbateľný:

$$\left|\Pr\left[D\left(g^x, g^t, g^{xt}\right) = 1\right] - \Pr\left[D\left(g^x, g^t, g^u\right) = 1\right]\right| = \left|\varepsilon(k) - \frac{1}{2}\right| \leq \mathrm{negl}(k)$$

- odtiaľ dostávame: $\varepsilon(k) \leq \frac{1}{2} + \mathrm{negl}(k)$

Poznámky k dôkazu

- alternatívny spôsob dokazovania – *game hopping*
- definujeme postupnosť hier (experimentov, ako napr. Ind^{cpa})
  - začíname s hrou podľa definície (G_0)
  - postupne upravujeme niektoré časti (voľba parametrov, výpočet a pod.)
  - končíme s hrou, v ktorej útočník nemôže vyhrať (G_n)
- ukážeme, že rozdiel pravdepodobností výhry medzi G_i a G_{i+1} je zanedbateľný, napr.
  - nerozlíšiteľnosť na základe nejakého problému
  - identické hry až na chybovú udalosť so zanedbateľnou pravdepodobnosťou
  - matematicky identické hry (len inak, ekvivalentne zapísané)
- V. Shoup: *Sequences of Games: A Tool for Taming Complexity in Security Proofs*, 2006

Definícia bezpečnosti – experiment (IND-CCA2)

- výpočtovo obmedzený (polynomiálny) útočník A
- útočník má prístup k verejnému kľúču a k dešifrovaciemu orákulu $\mathrm{Dec}_{\mathrm{sk}}(\cdot)$
- cieľ: rozlíšiť, z ktorého z dvoch otvorených textov vznikol šifrový text

$\mathbf{Ind}^{\mathbf{cca2}}_{A,\Pi}(\boldsymbol{k})$:

$(\mathrm{pk}, \mathrm{sk}) \leftarrow \mathrm{Gen}(1^k)$

$(m_0, m_1, s) \leftarrow A^{\mathrm{Dec}_{\mathrm{sk}}}(\mathrm{pk})$, kde $|m_0| = |m_1|$ a s reprezentuje stav výpočtu

$c \leftarrow \mathrm{Enc}_{\mathrm{pk}}(m_b)$, kde $b \in_R \{0, 1\}$

$b_A \leftarrow A^{\mathrm{Dec}}_{\mathrm{sk}}(s, c)$, pričom sa A nesmie pýtať priamo $\mathrm{Dec}_{\mathrm{sk}}(c)$

return $b_A \stackrel{?}{=} b$ (1 ak A určil b správne, inak 0)

Definícia bezpečnosti – IND-CCA2

Definícia. Šifrovacia schéma $\Pi = \langle \mathrm{Gen}, \mathrm{Enc}, \mathrm{Dec} \rangle$ je IND-CCA2 bezpečná, ak pre ľubovoľného PPT útočníka A existuje zanedbateľná funkcia $\mathrm{negl}(\cdot)$ taká, že platí:

$$\Pr\left[\mathrm{Ind}^{\mathrm{cca2}}_{A,\Pi}(k) = 1\right] \leq \frac{1}{2} + \mathrm{negl}(k).$$

– variant experimentu a definície IND-CCA1: A po získaní c nemá vôbec prístup k $\mathrm{Dec}_{\mathrm{sk}}$ orákulu

- A postupuje takto:
  1. zvolí ľubovoľné $m_0 \neq m_1 \in G$
  2. získa šifrový text $c = (r, s)$
  3. využije prístup k dešifrovaciemu orákulu: $\mathrm{Dec}_{\mathrm{sk}}(r, s \cdot m')$, kde $m' \in G$ je ľubovoľné
  4. získaný otvorený text $m_b \cdot m'$ použije na určenie b

Príklady IND-CCA2 bezpečných schém

- RSA-OAEP
  - model s náhodným orákulom
  - RSA predpoklad (nie tesná redukcia)
  - iné konštrukcie ako RSA použité s OAEP vyžadujú *partial-domain one-wayness*
- Cramer-Shoup
  - *štandardné* kryptografické predpoklady
  - DDH predpoklad
  - hašovacia funkcia s vlastnosťou *target collision resistance / universal one-wayness* (≈ hľadanie druhého vzoru vopred zvolenej správy pre rodinu hašovacích funkcií)
- ideálne modely (náhodné orákulá, ideálne šifry a pod.) vs. štandardné predpoklady
  - *neinštancovateľné* schémy

FO transformácia

- Fujisaki, Okamoto (1999)
- ako dosiahnuť IND-CCA2 bezpečnú schému skombinovaním slabších schém
  - hybridná konštrukcia (šifrovanie)
  - generická konštrukcia v modeli s náhodným orákulom
  - slabšia asymetrická schéma (v podstate OW-CPA, jednosmernosť)
  - γ-uniformita pre asymetrickú schému (pravd. ľub. dvojice OT, ŠT je $\leq \gamma$)
  - slabšia symetrická konštrukcia (Find-Guess $\sim$ IND, útočník bez šifr. orákula)
- viaceré vylepšenia a modifikácie neskôr
- redukcia nie je tesná

FO transformácia – ElGamalova schéma s OTP

- Gen – ako predtým, m je z priestoru správ symetrickej šifry
- H_1, H_2 náhodné orákulá s vhodnými definičnými obormi aj obormi hodnôt
- v ElGamalovej schéme šifrujeme náhodne zvolenú správu σ

$\mathrm{Enc}_{\mathrm{pk}}(m) \mapsto (r, s, c)$:

$\sigma \in_R G$

$t = H_2(\sigma, m)$

$(r, s) = \left(g^t, \sigma \cdot y^t\right)$

$c = m \oplus H_1(\sigma)$ resp: $E^{\mathrm{sym}}_{H_1(\sigma)}(m)$

$\mathrm{Dec}_{\mathrm{sk}}(r, s, c)$:

$\sigma' = s \cdot r^{-x}$

$m' = c \oplus H_1(\sigma')$ resp: $D^{\mathrm{sym}}_{H_1(\sigma')}(c)$

if $s = \sigma' \cdot y^{H_2(\sigma', m')}$: $m = m'$

else: $m = \perp$ (explicitné zamietnutie)

return m

Dôkazy v kryptológii – záver

- dokázateľná bezpečnosť $\not\Rightarrow$ bezpečná schéma v praxi
- dôkazy môžu byť chybné (pôvodný dôkaz pre RSA-OAEP)
- definície bezpečnosti nemusia zodpovedať praktickým potrebám
- inštancie schém nezodpovedajúce pravdepodobnostným odhadom (nie tesné redukcie)
- predpoklady bezpečnosti nemusia byť naplnené (náhodné orákulá)
- nevhodná implementácia (útoky postrannými kanálmi, ukladanie kľúčov, entropia, …)
- obídenie kryptografie (ľudský faktor)
- *napriek všetkému to nie je zbytočné*

Úlohy

1. Je DDH ťažký v grupe $(\mathbb{Z}_p^*, \cdot)$, kde p je prvočíslo?
2. Skonštruujme šifrovaciu schému z RSA: $c = (r^e \bmod n, \mathrm{lsb}(r) \oplus m)$, kde $m \in \{0, 1\}$ a $r \in_R \mathbb{Z}_n$. Zdôvodnite, že táto schéma je IND-CPA bezpečná. Aký predpoklad je potrebný? Je schéma aj IND-CCA2 bezpečná?
   - *poznámka:* modifikácia s $H(r)$ namiesto lsb bude IND-CPA bezpečná v modeli s náhodným orákulom